# THE USE OF PARTIAL UPDATE SCHEMES TO REDUCE INTER-CHANNEL COHERENCE IN ADAPTIVE STEREOPHONIC ACOUSTIC ECHO CANCELLATION

*Andy W. H. Khong and Patrick A. Naylor*

Department of Electrical and Electronic Engineering, Imperial College London, UK
E-mail: {andy.khong, p.naylor} @imperial.ac.uk

### ABSTRACT

The use of partial update adaptive filters for stereophonic acoustic echo cancellation is investigated. The MMax-NLMS algorithm is studied in this context and its inherent robustness to subsampling of the tap-input vector is demonstrated. It is proposed to employ the subsampling of the tap input vector, that is intrinsic to partial update schemes, to decorrelate the two tap-input vectors of the stereo adaptive structure thereby enhancing convergence. We investigate the trade-off between improvement in convergence due to decorrelation of the two tap-input vectors and degradation in convergence due to subsampling in the MMax-NLMS partial update scheme. The exclusive MMax-NLMS (XM-NLMS) is proposed which approximates the joint optimization of these factors and simulation results are presented.

## I. INTRODUCTION

Direct application of adaptive filters to the problem of stereophonic acoustic echo cancellation (SAEC) is known to be ineffective due to the high coherence between the two input signals. This has led to several approaches to the problem that involve techniques to decorrelate the two input signals using, for example, non-linear processing [1] or additive signals [2]. Furthermore, the computational complexity of stereophonic echo cancellers can be high because the number of taps can be large and also because the use of least-squares-based algorithms is often preferred in order to obtain sufficient levels of cancellation. Therefore, there exists a dual motivation to develop algorithms which have improved convergence performance due to reduction of interchannel coherence whilst maintaining computational complexity to be as low as possible for practical reasons.

In partial update adaptive filtering, the tap-input vector is subsampled so that only a subset of filter taps is updated at each iteration [3] [4]. The aim of this work is to investigate whether such subsampling can bring a reduction in the inter-channel coherence of the tap-input vectors that results in improved convergence.

The problem has been structured as a joint optimization of two scores - one describing the inter-channel coherence between the tap-input vectors and the other describing the 'closeness' of the tap selection to that of the MMax-NLMS scheme. In this context, the ideal tap selection is therefore one which selects the elements of the tap input vectors such that the inter-channel coherence is minimized whilst maximizing their $L_1$ norm.

A brief discussion of MMax-NLMS is presented in Section II. We shall look at the effect of decorrelation in Section III. Section IV presents XM-NLMS algorithm while Section V concludes the present work and discusses the ongoing research work.

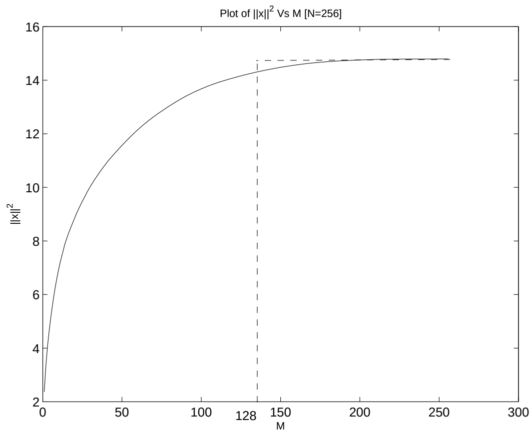


**Fig. 1**. Effect of subselecting update weights on squared Euclidean Norm

## II. MMax-NLMS ALGORITHM

In the MMax-NLMS algorithm [5], only the weights corresponding to the *M* largest tap inputs are updated at each iteration. It has been shown that, under specified conditions, the rate of convergence of MMax-NLMS depends on *M* whilst the same final misadjustment as the NLMS is achieved. To show the robustness of the MMax-NLMS, (1) and (2) establish the relationship between the convergence rate and the squared Euclidean norm of the tap-input vector $\mathbf{x}(n)$ for a single channel environment.

$$\|\varepsilon(n+1)\|^2 - \|\varepsilon(n)\|^2 = \mu^2\|\mathbf{x}(n)\|^2 e^2(n) - 2\mu e^2(n) \quad (1)$$

$$\|\varepsilon(n)\|^2 = \sum_{i=1}^{L-1}(h_i(n) - \tilde{h_i}(n))^2 \quad (2)$$

where $h_i(n)$ and $\tilde{h_i}(n)$ are the impulse response of the receiving room and the adaptive filter tap weights respectively. We also define, $\mu$, as the adaptive step size, $e(n)$ is the difference between desired signal and the output of the adaptive filter of length $L$.

Figure 1 shows the effect of sub-selection on $\|\mathbf{x}(n)\|^2$ for an FIR filter of length *L*=256. For every value of $M < L$, the squared Euclidean Norm of the *M* highest inputs is measured and plotted over 100 trials. The white noise input $\mathbf{x}(n)$ is of zero mean and unit variance. It can be seen that within the region $128 \leq M \leq 256$, there is only a modest reduction in $\|\mathbf{x}(n)\|^2$. In the region $M \leq 128$, there is an approximate exponential reduction in $\|\mathbf{x}(n)\|^2$. Using this relationship and (1), one can see that there is only a modest reduction in convergence rate when the filter is operating within the region ($0.5L \leq M \leq L$).

### III. THE EFFECT OF DECCORELATION IN SAEC

A serious problem encountered in stereophonic echo cancellers is that the echo canceller coefficients do not necessarily converge to the true impulse response of the echo path when full modelling of the transmission room exists (nonuniqueness problem) [1] [6]. In a practical case where the length of the filter is less than the impulse response of the transmission room, the problem of nonuniqueness is ameliorated to some degree. However, due to the strong coherence between the two channel input signals, the convergence performance is poor and misalignment is a significant problem. To overcome the misalignment problem, it is therefore desirable to reduce the coherence between the two input channel signals.

Although the tap selection concept underlying partial update algorithms, such as MMax-NLMS, is sometimes employed with the aim of complexity reduction, this is not the case in this work. Instead, we consider the use of tap selection to reduce the coherence of the two channel inputs. It is also known that if the two channels are highly coherent, the tap-input vectors are very similar. This will cause the MMax-NLMS algorithm to select and update the corresponding weights for the two filters. This does not achieve our desired effect of decorrelating the signals. We therefore introduce an alternative tap selection criterion controlled by two variables: magnitude weighting, to describe the 'closeness' of the tap selection to that of the MMax-NLMS scheme, and coherence weighting, to describe inter-channel coherence between the subsampled tap-input vectors respectively.

For the following description of XM-NLMS, we assume a standard FIR adaptive filter configuration. We define **A** and **C** being square matrices containing elements $a_{ij} = \|x\{\alpha_{ij}\}\|$ and $c_{ij} = coh\{\alpha_{ij}\}$ respectively where $\{\alpha_{ij}\}$ is a tap selection set with $i$ and $j$ representing the indices for each channel of the different combinations of selecting $M$ out of $L$ weights in each of the two filters $(i, j = 1, \ldots, {}^{L}C_M)$. The absolute sum of the selected tap inputs in a particular combination *i* and *j* for the two channels is defined as $\|x\{\alpha_{ij}\}\|$ while $coh\{\alpha_{ij}\}$ is the coherence averaged over frequency of the two tap-input vectors with $L-M$ unselected inputs in each channel set to zero. Elements $a_{ij}$ and $c_{ij}$ are each associated with a cost such that the least cost is allocated to combinations having the maximum magnitude in **A** and the minimum coherence in **C**. The update equation incorporating tap selection is given by:

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mathbf{G}(n)\mu\frac{\mathbf{x}(n)e(n)}{\|\mathbf{x}(n)\|^2} \tag{3}$$

where $\mathbf{G}(n) = diag\{g_{ij}(n)\}$ such that

$$g_{ij}(n) = \begin{cases} 1 & \text{if } i,j \in \{\alpha_{min}\} \\ 0 & \text{otherwise} \end{cases}$$

where $\{\alpha_{min}\}$ is the tap selection set for minimum cost and

$$\tilde{\mathbf{h}} = [\tilde{\mathbf{h}}_1^T \tilde{\mathbf{h}}_2^T]^T$$

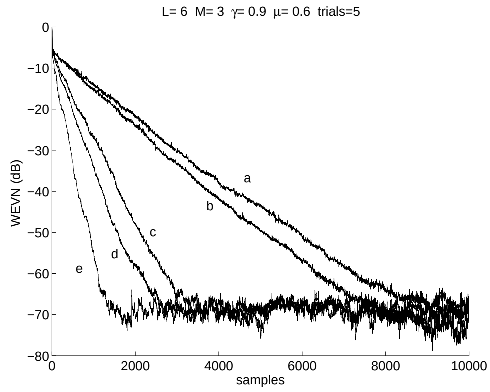


**Fig. 2**. WEVN for (a) MMax-NLMS and XM-NLMS $w_m = 1$, (b) NLMS (c) XM-NLMS $w_m = 0.9$, (d) XM-NLMS $w_m = 0.7$, (e) XM-NLMS $w_m = 0.1$. *L*=6, *M*=3, $\mu = 0.6$, $\gamma = 0.9$.

represents the concatenated weight vector of the two channel filters while $^T$ represents vector transposition. Magnitude Weighting, $w_m$, and coherence weighting, $w_c = 1 - w_m$, determine the 'closeness' of the tap selection to that of the MMax-NLMS scheme. A magnitude weighting of 1 corresponds to selecting coefficients for updating based on the MMax-NLMS algorithm only.

Figure 2 shows the weight error vector norm defined as [1]

$$WEVN = \frac{\|\mathbf{h} - \tilde{\mathbf{h}}\|}{\|\mathbf{h}\|} \tag{4}$$

for different values of magnitude weighting ($w_m = 0.1, 0.7, 0.9, 1.0$). In this simulation, the two channel inputs are zero mean and unit variance noise sequences. The coherence between the two channel inputs is controlled by $\gamma$ ($0 \leq \gamma \leq 1$), where $\gamma$=0 represents independent signals and $\gamma$=1 implies the two channel inputs being identical to one another. In this simulation, $\gamma$=0.9 is used to reflect the high coherence of the two channel weight input vectors in practical applications. The adaptive filters have 6 taps per channel and for every iteration, 3 taps are updated ($L = 6, M = 3$). For clarity, WEVN for only one of the two channels is plotted for each case of $w_m$.

The simulation result shows that $w_m$=1 coincides with MMax-NLMS where performance is close to that of the full update NLMS. The highest convergence rate can be seen when $w_m$=0.1 ($w_c$=0.9) where there is a high emphasis on selecting the exclusive set of weights for updating. This is similar to finding (out of ${}^{L}C_M$ combinations) the exclusive set of weights for the two filters such that the absolute sum of the subsampled tap-input vector is maximized.

### IV. XM-NLMS ALGORITHM

Consider initially a pair of adaptive filters (each of length $L$=4) with tap input vectors $\mathbf{q} = [q_1, q_2, q_3, q_4]^T$ and $\mathbf{r} = [r_1, r_2, r_3, r_4]^T$. Let **d** be the magnitude vector difference between **q** and **r** such that $\mathbf{d} = |\mathbf{q}| - |\mathbf{r}|$,

$$\begin{pmatrix} |q_1| \\ |q_2| \\ |q_3| \\ |q_4| \end{pmatrix} - \begin{pmatrix} |r_1| \\ |r_2| \\ |r_3| \\ |r_4| \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \end{pmatrix}. \tag{5}$$

Our objective is to select $M$ out of $L$ taps for updating for which the corresponding (subsampled) tap-input vectors have the maximum absolute sum, so as to approximate MMax-NLMS as closely as possible, but also have the minimum inter-channel coherence. Whereas in principle an exhaustive search of all possible tap selection sets could be made for small $L$, the XM-NLMS algorithm finds an approximation to the optimum tap selection by constraining the search to tap selections that are exclusive between the two channels so as to minimize the inter-channel coherence. These exclusive sets can be pre-determined for any $L$ and $M$. Within this constrained search space, the tap selection with maximum absolute sum can be found efficiently by sorting **d**. With reference to (5), assuming $d_1 > d_2 > d_3 > d_4$, then we have

$$|q_1| + |q_2| + |r_3| + |r_4| > |q_3| + |q_4| + |r_1| + |r_2| \qquad (6)$$

since $d_1 + d_2 > d_3 + d_4$. Thus tap weights corresponding to input $q_1, q_2, r_3$ and $r_4$ should be selected for updating. It is worthwhile noting that it is irrelevant to consider combinations $\{d_1, d_3\}$, $\{d_1, d_4\}$, $\{d_2, d_3\}$, $\{d_2, d_4\}$ or $\{d_3, d_4\}$ since the corresponding tap selections in these combinations would give a smaller magnitude sum than the tap selection in combination $\{d_1, d_2\}$. This approach allows us to eliminate $({}^{L}C_M - 1)$ possible combinations.

We now consider two filters each of arbitrary length $L$ where we update $M$ taps from each filter. A typical case is when $M = 0.5L$. We can extend (5) as,

$$\begin{pmatrix} |q_1| \\ |q_2| \\ . \\ . \\ . \\ |q_L| \end{pmatrix} - \begin{pmatrix} |r_1| \\ |r_2| \\ . \\ . \\ . \\ |r_L| \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ . \\ . \\ . \\ d_L \end{pmatrix}. \qquad (7)$$

As before, if $d_1 > \ldots > d_M > \ldots > d_L$, we now select taps corresponding to inputs $\{q_1, ..., q_M\}$ and $\{r_{M+1}, ...r_L\}$. Thus the two tap input vectors of length $L$ are ordered in descending order of difference in absolute value **d** and the first $M$ taps of the ordered vectors are selected for updating. Equation (8) shows the updating equation of the XM-NLMS algorithm.

$$\tilde{\mathbf{h}}(n+1) = \tilde{\mathbf{h}}(n) + \mathbf{G}(n)\mu\frac{\mathbf{x}(n)e(n)}{\|\mathbf{x}(n)\|^2} \qquad (8)$$

where $\mathbf{G}(n) = diag\{g_{ij}(n)\}$ such that

$$g_{ij}(n) = \begin{cases} 1 & \text{if } i,j \in \{\beta_{min}\} \\ 0 & \text{otherwise} \end{cases}$$

where $\{\beta_{min}\}$ is the tap selection set for minimum cost subject to the above constraint that the selection is exclusive between the two channels.

As can be seen from (8), XM-NLMS can be implemented using a preprocessor block located at the input of the adaptive filter pair. The purpose of this preprocessor block is to perform the exclusive MMax tap selection. This structure separates the tap selection from adaptive filter block and therefore enables us to employ alternative adaptive filter blocks so as to form, for example, the exclusive MMax recursive least squares algorithm (XM-RLS).

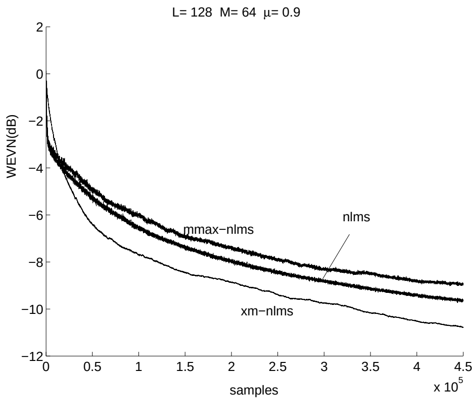


**Fig. 3**. WEVN for MMax-NLMS, NLMS and XM-NLMS ($L$=128, $M$=64, $\mu = 0.9$)

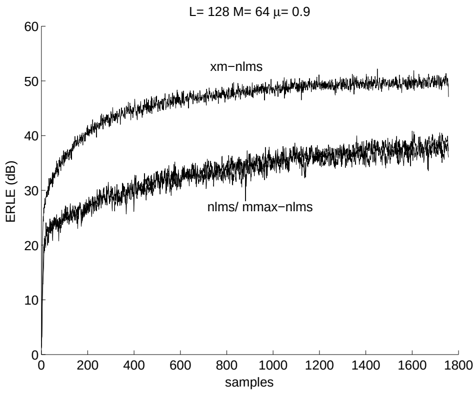


**Fig. 4**. ERLE for MMax-NLMS, NLMS and XM-NLMS ($L$=128, $M$=64, $\mu = 0.9$)

## V. RESULTS

The XM-NLMS algorithm was applied in an SAEC experiment using Gaussian noise with zero mean and unit variance as the input sequence. The transmission room microphone signals were obtained by convolving the source with two white Gaussian impulse responses $g_1(n)$ and $g_2(n)$ of length 800 with $\gamma$=0.9. The desired response in the receiving room was obtained by summing the two convolutions $(h_1 * x_1)$ and $(h_2 * x_2)$ where $h_1$ and $h_2$ represent the receiving room's impulse responses. The method of images [8] was used to generate $h_1$ and $h_2$ of length 800 which were subsequently truncated to 128. The adaptive filters were of length $L$=128, $M$=64 for the XM-NLMS and MMax-NLMS. Figure 3 shows the WEVN plot for NLMS, XM-NLMS and MMax-NLMS algorithm. Figure 4 shows the corresponding segmental echo return loss enhancement (ERLE). For all our experiments, we have used a step size $\mu = 0.9$.

As expected, due to the decorrelating property, we can see from Fig. 3 that XM-NLMS converges at a faster rate. As pointed out in [1], the non-linear preprocessor is ineffective in conjunction with NLMS. Thus, we have excluded the NLMS with non-linear preprocessor in this experiment. The results of Fig. 5 show the convergence performance of XM-RLS using a preprocessor block

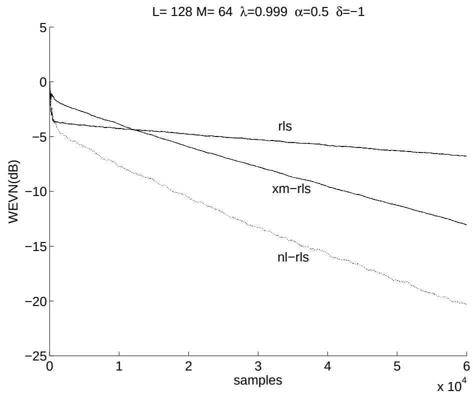


**Fig. 5**. WEVN for RLS, XM-RLS and NL-RLS for zero mean unit variance $g_1$ and $g_2$. ($L$=128, $M$=64, λ=0.999, $\delta$=-1, $\alpha$=0.5)

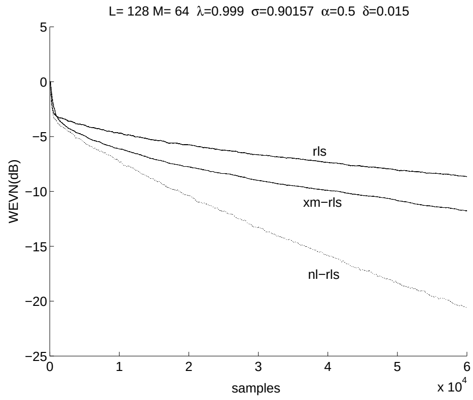


**Fig. 6**. WEVN for RLS, XM-RLS and NL-RLS. $g_1$ and $g_2$ generated using method of images. ($L$=128, $M$=64, λ=0.999, $\delta$=0.015, $\alpha$=0.5)

explained in the previous Section. The input signal is white Gaussian noise with zero mean and unit variance. We have used the same impulse responses as before for both receiving and transmission rooms. The RLS parameters are as follows: forgetting factor $\lambda$=0.999 and regularization parameter $\delta$=-1 [7]. In NL-RLS the non-linearity constant $\alpha = 0.5$.

Figure 6 shows an additional result for the case when both the transmission and receiving rooms' impulse responses were generated using the method of images [8]. In this simulation, two microphones are placed one meter apart in the center of each room (size 3x4x5 metres). The source is then positioned such that it is one meter away from each of the microphones in the transmission room. The impulse responses $g_1$, $g_2$, $h_1$ and $h_2$ are each of length 800. As before, we have used the following RLS parameters; forgetting factor $\lambda$=0.999 and regularization parameter $\delta$=0.015. The convergence behavior is plotted in Fig. 6 for XM-RLS, RLS and NL-RLS (over an average of 3 trials).

We can see from Fig 5 and Fig 6 that XM-RLS achieves an improved rate of convergence compared to RLS. This is again due to the intrinsic decorrelation property of the exclusive MMax preprocessor block acting on the input signals.

## V. CONCLUSION

In this paper, we have discussed the robustness of MMax-NLMS to subsampling of the tap-input vector for a general single channel case. We have also shown in SAEC how the two tap-input vectors are effectively decorrelated using the MMax subsampling procedure with a resulting improvement in convergence in WEVN. We have tested the approach using room impulse responses generated randomly and also realistically using room modelling with higher order. An efficient technique has been proposed for determining a tap selection set that gives an approximate joint optimization of maximum absolute sum of the subsampled tap-input vectors and minimum inter-channel coherence. Both NLMS and RLS adaptive filters show improved convergence performance in SAEC when used in conjunction with the partial updating scheme. Further testing is underway with speech signals.

## REFERENCES


[1] J. Benesty, D. R. Morgan and M. M. Sondhi, "A Better Understanding and an Improved Solution to the Specific Problems of Stereophonic Acoustic Echo Cancellation", *IEEE Trans. Speech and Audio Processing*, vol. 6, no. 2, Pages. 156-165, 1998.

[2] M. M. Sondhi, D. R. Morgan and J. L. Hall, "Stereophonic Acoustic Echo Cancellation- An Overview of the Fundamental Problem", *IEEE Signal Processing Letters*, vol. 2, no. 8, pp. 148-151, 1995.

[3] K. Dogancay and O. Tanrikulu, "Adaptive Filtering Algorithms With Selective Partial Updates", *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, Volume: 48 Issue: 8 , Aug 2001. Pages: 762 -769.

[4] P. A. Naylor and W. Sherliker, "A Short-Sort M-Max NLMS Partial Update Adaptive Filter with Applications to Echo Cancellation", *IEEE International Conference on Acoustics, Speech and Signal Processing*,2003.

[5] T. Aboulnasr and K. Mayyas, "Complexity Reduction of the NLMS Algorithm via Selective Coefficient Update", *IEEE Trans on Signal Processing*, vol. 47, no. 5, Pages. 1421-1424, 1999.

[6] S. Shimauchi and S. Makino, "Stereo Projection Echo Canceller With True Echo Path Estimation", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1995. ICASSP-95 , Volume: 5 , 9-12 May 1995, Pages: 3059 -3060

[7] S. Haykin, "Adaptive Filter Theory 4th Edition," *Prentice Hall*, 2002

[8] J. B. Allen and D. A. Berkley, "Image Method for Efficiently Simulating Small-Room Acoustics," *J. Acoust. Soc. Amer.*, vol.65, pp. 943-950, Apr. 1979.